Detecting Object Motion with Passive RFID Technology

Siddika Parlak, *Student Member, IEEE,* Ivan Marsic, *Member, IEEE*

***Abstract*—In this paper we explore the problem of detecting and classifying object motion using passive RFID technology. We extract the useful features of the received signal strength sequence for classifying the object motion as *moving linearly, moving randomly or standing still*. We experimented with several machine-learning algorithms, from non-temporal classification methods to generative and discriminative sequence labeling algorithms. We also compared the performance of single stage and cascaded classification on object motion monitoring. Finally, we present an analysis of the individual contribution of features. We observed that …. Gives the best results. Experiments, performed both in ideal conditions and more cluttered conditions, showed that object motion detection is feasible in environments with multiple tags and human movement.**

***Index Terms*—About four key words or phrases in alphabetical order, separated by commas. For a list of suggested keywords, send a blank e-mail to keywords@ieee.org or visit** http://www.ieee.org/organizations/pubs/ani_prod/keywrd98.txt

I. Introduction

Mobility is an important component of context-aware systems in a variety of domains. As an example, localization accuracy can be improved by ignoring the jumps in the signal strength when a user is stationary [1]. In [2], frequency of scans for localization is adjusted based on mobility information, because active scans are required only when the user is in motion. In healthcare, the activity level of a person can be estimated and automatic assistance can be provided for healthy living [1]. Equipment usage on patients can be monitored in a hospital. In a home setting, moving items (due to interaction by a person) can be identified and daily activities of elderly people can be recognized [3]. Although, identification with wearable sensors is the most widely used technique for this kind of applications, wearable sensors may not be appropriate for some settings, or can be aided with object motion detection (e.g. when user holds the object with the wrong hand).

In this work, we explore the performance of passive UHF RFID technology for detecting and classifying small-scale object motion (e.g. within a room). We prefer passive RFID as the sensing technology because passive tags enable even very small and low-cost objects to be identified. To date, monitoring of object/human mobility is studied using several sensing mechanisms such as GSM, GPS, Wi-Fi and active RFID; however passive RFID has received limited attention. All of the mentioned sensing mechanisms, except passive RFID, require an active transmitter to be attached on the object, which is not appropriate for small items. Motion status is inferred based on the strength of the received signal generated at the transmitter. However a passive RFID tag cannot generate signal but scatters the signal coming from the reader. This results in shorter read ranges compared to Wi-Fi and more sensitivity to other factors such as orientation and cluttered environments.

In this work, we classify the motion type of an object as: *still* (S), *moving linearly* (L) or *moving randomly* (R), instead of mere classification into "still" or "moving". Moving linearly represents changing position with few orientation changes and small variations in movement. Moving randomly represents experiencing many small movements and orientation changes without actual position change. Identifying the movement type facilitates detecting correct usage status, because random movement is a stronger indicative of item usage. Relocating an item does not necessarily shows that it is in-use. Consider an emergency room where a nurse fetches a thermometer to measure the patient's temperature. Since the patient has a mask on his face, the nurse cannot obtain the temperature at the moment so she leaves the thermometer on the patient bed. After several minutes, she takes the thermometer again and measures the temperature. With the assumption that any interaction indicates usage, the first interaction in this example is clearly a false alarm.

Our methodology is based on processing the Received Signal Strength (RSS) sequence within a sliding window to extract the useful features for motion classification. We focused on the fluctuations in the signal strength as the representative statistical features. Several of them are used, such as standard deviation, Euclidean distance, and delta mean. Moreover we propose and discuss interpolation as a preprocessing method.

The feature sequence is labeled using non-temporal (Decision Tree, Boosting, Support Vector Machine), generative-temporal (Hidden Markov Model) and discriminative temporal classifiers (Conditional Random Field). For non-temporal classifiers, both single-stage

Parts of this study were presented in conference [parlak10]
Siddika Parlak and Ivan Marsic are with the Department of electrical Engineering, Rutgers University, Piscataway, 08854, NJ, USA. E-mail: parlak@rci.rutgers.edu, marsic@ece.rutgers.edu.

multiclass classification and 2-stage binary classification approaches are evaluated. We conducted the experiments in several settings, such as multiple tags and human motion in the environment.

A. Motivation: Trauma Resuscitation

Our target application is in healthcare domain, specifically Trauma Resuscitation[1]. Object motion gives clues about object usage, which can be used to infer performed medical activities, such as intubation or blood pressure measurement. Previous work has shown that, context-aware and decision-support systems can be designed for trauma resuscitation by recognizing performed tasks based on the used medical equipment [4].

Consider a usage scenario for a blood pressure (BP) cuff, which is initially at rest (standing still). When a blood pressure is needed, a team member fetches the BP cuff, brings it near patient (position change – linearly moving) and places on patient's arm (randomly moving). Finally, being placed, the object is at rest again (standing still). Note that, relocating a BP cuff does not justify usage itself because some random movements are expected as well. We deduced relying on a large number of similar observations that, human interactions with objects translate into three types of object motions: (i) standing still; (ii) moving linearly; and (iii) moving randomly.

Because the items for tagging in the trauma bay vary in amount and size, passive UHF RFID tags are appropriate. Wearable readers are not convenient because urgency of the situation during resuscitation events may make the workers forget or ignore wearing readers. Moreover, trauma resuscitation is a dynamic work environment with team members constantly joining and leaving the room. It is therefore difficult to equip the entire team with wearable readers and ensure that all active members are wearing them throughout the process. Because of these factors, fixed readers are more appropriate than wearable readers.

A major concern for RFID deployment in medical domains is the possible electromagnetic interference by RFID on critical care equipment. In [vanderTogt08] passive RFID system was reported to induce incidents in 63% of the trials. The median distance between the reader and the medical device in all interference incidents was 30 cm. Another study

Although trauma resuscitation is a typical application for small-scale object motion detection, our work and results are not limited to the described application.

II. Related Work

A. Mobility Monitoring

In, [5] a traveler's transportation mode (*walking, driving* or *taking a bus*) was inferred, and his most likely route was predicted from GPS data using Dynamic Bayesian Networks. In [6], several machine learning algorithms, such as naïve Bayes, SVM and boosting, were used to infer user mobility as *stationary, walking* or *driving* from GSM data. Both studies base their algorithms on the coarse-grained location data, which is not applicable for detecting motion in a typical room-sized area, such as a trauma bay.

It is possible to obtain finer grained information using WLAN RSSI. A user's motion mode was inferred as *moving* or *standing still* in [7] using signal strength with 87% accuracy. The initial decision, based on the variance of the RSSI, was smoothed with a 2-state HMM. The ComPoScan [2] system incorporates an HMM-based motion detector for switching between monitor sniffing and active scanning to ensure the quality of communication while performing localization. Using spectral features of the WLAN RSSI signal, *moving* vs. *still* classification was performed in a diverse set of conditions in [1], with an average accuracy of 94%. A method is described in [8] for detecting intra-room mobility of users. The algorithm compares a set of signal strength descriptors (differences in mean, standard deviation and histogram comparison) with the previously estimated thresholds. Extensive experiments showed that their method is robust to several experimental settings, such as a conference room, a storage room or movement of a person in the environment. Although [8] and [2] are closest to our work in terms of granulation, mere classification of moving-not moving is not sufficient for our case. In general, WLAN technology is not appropriate for tracking small objects due to the aforementioned size limitations. In most of the methods highlighted above, either a temporal classifier based on signal strength variance, or a non-temporal classifier based on a richer set of features were used. However we will show in experiments section that, they are not sufficient to detect motion with passive RFID technology.

In [9], a human motion detection system based on passive RFID technology is introduced. Their method relies on HMMs and relative variance of response rate as observations. A change point detection algorithm is also integrated to adjust to different motion patterns. Different from [9], our aim is to detect mobility and usage of small items. To accomplish this, we label the motion mode of an object as: *still*, *moving linearly* or *moving randomly*. In addition we run experiments in challenging conditions such as high tag population and people movement in the environment.

B. Mobility Detection with RFID

An unobtrusive method of RFID tag motion detection with fixed readers was proposed in [10]. The algorithm makes use of the statistics computed within a sliding window and declares an interaction if the statistics exhibit a significant change. Unobtrusive detection of RFID tag motion was a part of the SixthSense platform as well [11]. They proposed an algorithm based on the difference between 10^{th} and 90^{th} percentiles of the received signal strength, again computed over a sliding window. Although [10] and [11] are close to our work in terms of sensing technology, they declare just the starting point of an

[1] Trauma resuscitation consists of a series of tasks performed to identify and immediately treat life-threatening injuries.

interaction, do not monitor motion continuously.

C. Object Detection with RFID (with near field readers, through identification)

Detecting used objects has been shown to be a fruitful approach for high-level activity recognition. The idea was applied to Activities of Daily Living (ADL) domain using wearable RFID readers and passive tags in [3] and [12], and to medical domain in [13]. In the mentioned studies, main focus was on the recognition of activities with the assumption that the sensor data includes no error. Reliability of several sensors was compared in ADL recognition in [14]. RFID was reported to be the worst performer because of a number of reasons such as indetectability of metallic objects and the inability of the reader to detect the object because of high tag-reader distance.

III. Sensors for Object Mobility Detection

This section gives an overview of RFID technology first and explains why we preferred RFID next.

A. RFID Overview

An RFID system consists of two basic components: a reader (interrogator) and a tag (transponder). A tag can be passive or active depending on whether it includes a battery. In a passive UHF RFID system, the reader generates an electromagnetic signal that provides energy for the tag to get activated and to transmit its unique ID (or perform other simple actions). The tag reflects part of the received signal while changing the impedance of its antenna. The reader decodes the ID of the tag by analyzing the reflected signal pattern. This process is known as *backscatter modulation* [15]. The energy received by the reader (Received Signal Strength Indication - RSSI) has an inverse relation with the distance traveled. Therefore, RFID can be exploited as a localization technology in addition to identification.

However, passive UHF RFID is designed for identification and there are challenges when it is used for other purposes, such as localization and motion detection. For maximum efficiency, polarization of the tag and reader antennas must be matched, which depends on the orientation of the tag. Inverse square law between tag-reader distance and RSSI gets highly complicated in a realistic environment because of fading, absorption, multipath and occlusions. High tag population and human motion are significant sources of noise in small environments, such as a typical room. In addition, correct measurement of the received signal strength at the reader is difficult due to poor isolation between transmit and receive channels. Resulting RSSI can be very noisy even when both the reader and antenna are static [16], [17].

B. Sensor Selection – why passive RFID?

Motion can be monitored using several sensing mechanisms, such as GPS, GSM, WLAN, accelerometers, RFID and cameras. Among these, passive RFID tags and cameras are the only alternatives for monitoring small and inexpensive objects. GPS and GSM are not applicable for detecting intra-room mobility. WLAN, accelerometers and active RFID tags are not convenient for tracking small and inexpensive objects (e.g. a toothbrush or an endotracheal tube[2]) because the sensor to be attached on the object is active – has its own energy source. A battery increases not only the cost, but also the sensor size.

Vision based methods offer a number of advantages over passive RFID. Because a considerable amount of research has been done in vision-based object detection and tracking, processing techniques are mature. Vision is not sensitive to object material, washing, sterilization etc. (whereas passive RFID has performance issues near liquids and metals). There is neither overhead of sensors on objects nor human effort required to place sensors.

On the other hand, RFID is advantageous when the target object is very small (for vision algorithms). RFID does not require line-of-sight and is able too work in case of human occlusion. However, vision algorithms cannot work in case of occlusion. RFID offers better scaling to a large number of unknown objects because no retraining is required when increasing the number of recognizable objects. RFID is better when parallel processing is required for a large number of objects. If privacy is a significant concern, RFID is the only choice.

In this work, our aim is to explore the motion detection performance of passive RFID alone. Depending on the particular application, either RFID or vision or hybrid of them can be optimum. Our results give insights about limits of object motion detectability with passive RFID and possible hybrids with vision.

IV. Experimental Setup

1) Environmental Setting

The experimental equipment consisted of an Alien RFID reader (Model: ALR-9900 (Four Antenna / Gen 2 / 902-928 MHz)), circularly polarized antennae (ALR-9611-CR) and passive RFID tags (Squiggle ALN-9540). Reader settings are summarized in Table I (See [21] for detailed explanations). Although only one reader is used in our experiments, we preferred Dense Reader Mode (DRM) for scalability of our results to larger deployments. Moreover, it has been shown in [17] that, DRM is the best performing mode when tag-reader distance is higher than 1.5 meters.

The experiments were conducted in a laboratory, which includes many objects that cause multipath and other adverse conditions for RF propagation. Environmental setting is depicted in Figure 1. One of the antennas was ceiling-mounted (2.7 meters above ground), facing the center of a plastic table of size 0.76x1.27x0.76 meters. Remaining two antennas were mounted to perpendicular sidewalls 2 meters above the ground, facing the table with an angle of 60 degrees with the x-axis. In this way, our aim was to detect movements in all three dimensions. Note that the environmental setting matches a

[2] A medical item that is inserted into the trachea for the primary purpose of establishing and maintaining an unobstructed airway

typical trauma resuscitation room. The plastic table mimics a patient bed. Considering that the trauma team gathers around the patient during trauma, ceiling mounted antenna is crucial in our design.

2) Data Collection

To create the dataset, an RFID tagged cartoon box (6″x4.5″x3″) was interacted with in the following ways:

- Not interacting at all (object standing still)
- Holding the box while walking around a table and from sidewalls to the table with an approximate speed of 40 inches/second (object moving linearly)

TABLE I
READER SETTINGS

Configuration	Definition
RF Modulation = Dense Reader Mode	This mode prevents interference among readers when many of them are operating in close proximity.
Persist time = -1	The reader is configured to perform a fresh scan at each interrogation cycle – previous tag list is not kept in memory.
RF Attenuation = 0	No RF attenuation is applied – the reader outputs 1 watt of RF power.
Acquire Mode = Inventory	The reader performs a full inventory of multiple tags (inventory mode), not a fast search for single tag (global scroll mode).
Auto Mode = ON	The reader operates on its own. An application on a host computer is set up to listen for notification messages from the reader containing any tag data that it has read.
Antennas = 0 1 2	The reader transmits through antennas 0, 1 and 2 in a round robin fashion.

- Standing at the same position and wiggling the box: rotating and occluding by fingers (object moving randomly at the same position).

Duration of each interaction type was limited to 20 seconds. The three interactions were performed continuously in various orderings to obtain an interaction sequence (event) of length 60 seconds. The dataset consists of 120 events performed in the following settings:

- One tag and no movement in the environment (60 events)
- One tag and a person moving in the environment—walking around the table and from sidewalls to the table with an approximate speed of 40 inches/second (30 events)
- 10 tags and no movement in the environment – 10 tags uniformly scattered on the table (30 events)

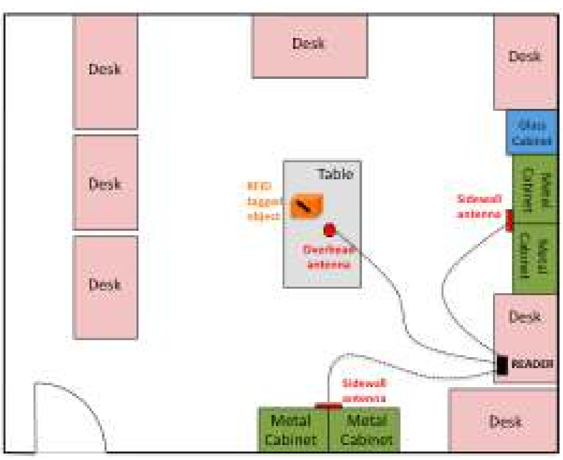

Fig. 1. The experimental setup.

Interrogation is performed through one antenna at a time in a round robin fashion. Because each antenna is visited for 0.5 seconds, duration of a complete cycle is 1.5 seconds.

V. METHOD

A. Feature Analysis

We extract the motion-related features from the RSS sequence using a fixed-length sliding window. At each time instant, the RSS values within the window are summarized with statistical features and the window is shifted for the next feature vector calculation.

1) Windowing

Several issues must be considered when determining the size of the sliding window. A longer window is useful in case of noisy observations because it provides smoother estimates. However short-duration movements might be missed with a long window. In addition, the latency of the system is proportional with the window size. For instance a 10 second window translates into a latency of 5 seconds, which might be long for many applications. Decision frequency of the system is related to the size of shift. A shift length of 2 seconds means that the system gives a decision once in 2 seconds.

In this work, we experimented with fixed-length windows of 1.5, 2, 3, 5 and 10 seconds. 1.5-second window is selected as the shortest because we have three antennas each with a scanning time of 0.5 seconds. 10-second window is selected as the longest because the latency and smoothing introduced by a longer window is not acceptable for many applications, such as trauma resuscitation (Section I.A). Various shift sizes were tested to analyze its effect on the performance and find the optimum one.

2) Interpolation

In our experiments, tag readings do not occur at regular intervals due to three factors:

- Reader is set to autonomous mode, tag readings may not arrive to the listener at regular intervals.

- Antennas are visited in a round-robin fashion. When the reader is transmitting from an antenna, data arrives only from that antenna.
- When multiple tags exist in the environment, read rate per tag reduces because of the collision detection mechanism.

Effect of these factors on the RSSI sequence is depicted in Figure 2. To reduce the effect of the irregularity in time axis, we interpolate the RSS values in a window. Number of interpolation points is determined based on the window length. (Such that 1 sample is captured in every 20 msecs.)

3) Features

For motion detection applications, features are selected to focus on the “changes” in the signal strength. Below, we present a list of coefficients for motion sensing.

Let the RSS values are captured with *n* antennas and the data points from antenna *i* within sliding window *t* are:

$$\mathbf{x}_i^t = \{x_{i1}^t, x_{i2}^t, \ldots, x_{in}^t\}$$

- Standard Deviation: High fluctuations in the RSS signal are indicators of movement. When the object is standing still, low standard deviation is expected [king08, li10].

$$std_i^t = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(x_{ij}^t - \bar{\mathbf{x}}_i^t)^2}$$

- Relative Variance: Using relative variance instead of variance (or standard deviation) might be useful when variance of signal increases with the increasing mean [li10].

$$relVar_i^t = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{x_{ij}^t - \bar{\mathbf{x}}_i^t}{\bar{\mathbf{x}}_i^t}\right)^2$$

- Euclidean Distance: When the object is standing still, RSSI values captured at close time instants are expected to be close to each other [muthu07].

$$euc_i^t = \frac{1}{n-C}\sum_{j=1}^{n-C}|x_{ij}^t - x_{ij+C}^t|$$

  To handle the irregularity in time axis and multiple tags, the RSS values are interpolated before calculating the Euclidean distance.

- Delta Mean: Delta mean is expected to be small for stationary objects and largest for linearly moving objects.

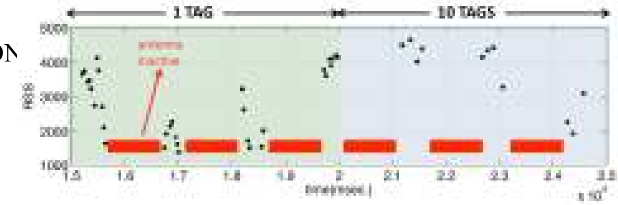

Fig. 2. A 10-second RSS capture from one of the antennas. One tag exists in the environment during the first 5 seconds and 10 tags exist in the environment during the last 5 seconds. Blocked regions show the intervals when the antenna is inactive (reader is transmitting through the other antennas.)

$$dmean_i^t = |(\bar{\mathbf{x}}_i^t) - (\bar{\mathbf{x}}_i^{t-1})|$$

- Minimum Range:

$$minR = \min_j(\max(\mathbf{x_j}) - \min(\mathbf{x_j}))$$

Figure 3 shows the variation of feature values for different movement types in different settings. The data points were obtained by averaging the feature values of three antennas.

For almost all feature types and settings, the feature value is smallest when the object is standing still and largest when the object is moving linearly. Comparing standard deviation and relative variance (Figures 3-a and 3-b), standard deviation offers better discrimination of standing still because data points have more intra-class similarity and inter-class dissimilarity. In addition, relative variance values in multiple tags setting are much lower for all motion types. Relative variance of a stationary object in ideal setting is higher than a linearly moving object in a multiple tag setting. This situation may cause problems if we use a parametric model for classification because parameters of ideal and multiple tag scenarios will not match.

Among the Euclidean distance descriptors (Figures 3-c, 3-d and 3-e), Euclidean distance with C=1 provides better discrimination, especially for still class. Separation of data points is even higher than the standard deviation feature (Figure 3-a). Comparing Figures 3-c-d-e, it is possible to deduce that as C gets higher discrimination gets worse.

Although delta mean can provide good discrimination in the ideal setting, this is not the case for multiple tag and movement settings.

The features are computed for each antenna separately and concatenated to obtain the feature vector. We preferred concatenation instead of averaging based on our preliminary experiments. Although moving-nonmoving classification was not affected by averaging, random-linear movement discrimination was worse. We observed that, linear movement is distinguished with a larger fluctuation (usually) coming from one antenna (possibly the one scanning in the direction of walking). When the average is taken, this fluctuation has a lower effect and the movement type cannot be detected correctly. Therefore, in case of RFID, it is not enough to simply take the average of multiple receivers. Orientation and placement are important factors, which will be discussed in detail in experimental results section. If data from an antenna is completely missing within a sliding window, knn imputation is applied.

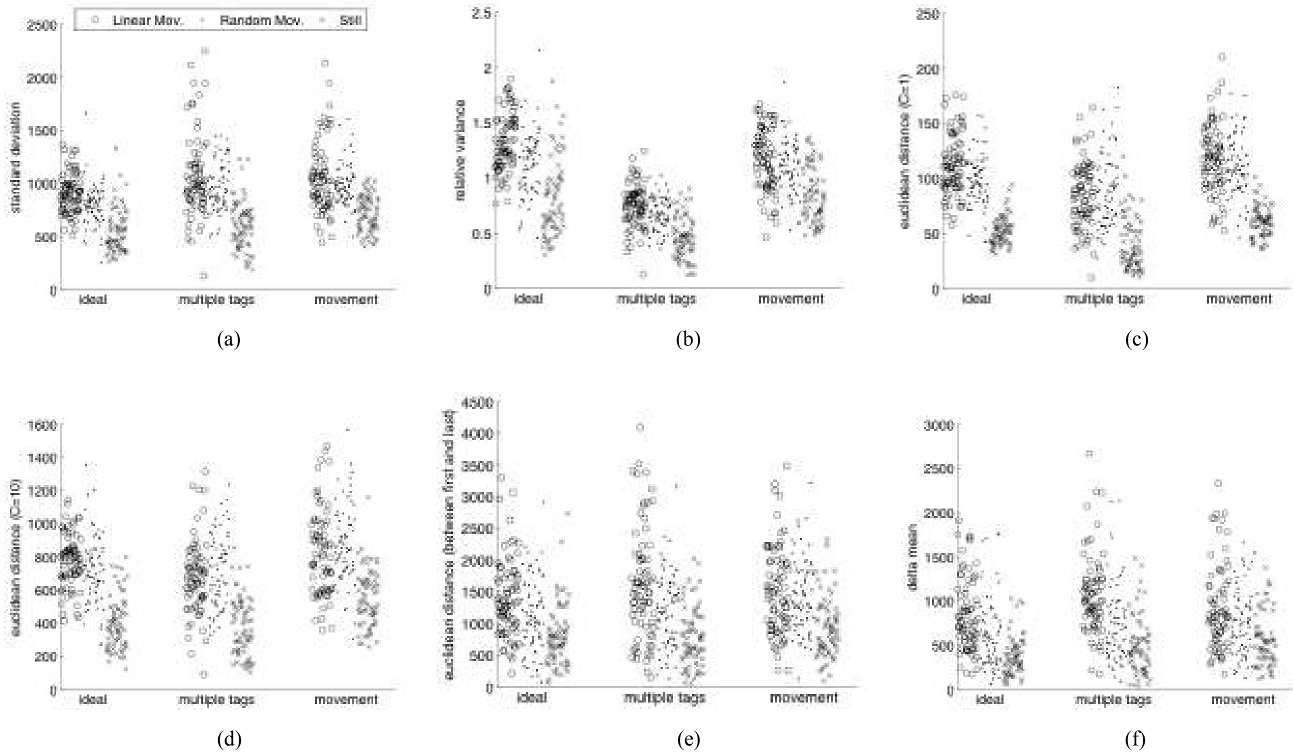

Figure 3. Various type of feature values obtained in different environments. Ideal: 1 tag, no movement. Multiple tags: 10 tags, no movement. Movement: 1 tag, human movement. (a) Standard deviation (b) Relative standard deviation (c) Euclidean distance (C=1) (d) Euclidean distance (C=10) (e) Euclidean distance (difference between first and last values in a window) (f) Delta mean.

B. Classification

1) Non-Temporal classification

The sliding window technique captures the temporal relations in the sequence of observed movements and allows non-temporal learning algorithms to be used on sequential data. (It does not capture temporal relations in the hidden state sequence.) To classify motions, we used SVM, Decision Tree and Boosting.

We built all non-temporal models with the WEKA Machine Learning Toolkit [18]. The SVM model is trained with a polynomial kernel. For the decision tree a C4.5 type tree is built and for Boosting, Adaboost.M1 method is used with Decision Stumps as base classifiers. Preliminary experiments showed that the label sequence obtained with the non-temporal classifiers included many spurious transitions, which correspond to short-term, frequently changing motions. An HMM-based post processing was applied on the classifier output to make the label sequence smoother. To estimate the observation probabilities of the smoothing HMM, training data is labeled with the SVM-based model and predicted labels are compared with the true labels.

2) Temporal Classification with Generative Models

Inferring the motion status over time is a sequential learning problem. An HMM is a probabilistic model, which takes the temporal (or spatial) information into consideration. HMMs have been successfully applied to many kinds of sequential learning problems in the literature, such as speech recognition and activity recognition. In this work, we built two types of HMMs: in the first one, named as the “uni-state representation”, each motion type is represented with a single state. In the second HMM, named as the “bi-state representation”, each type of motion is represented with a two-state left-to-right sub-HMM for modeling the motion transitions properly (Figure 4). *Main* is the core sub-state, representing the behavior of the state to which it belongs. *Exit* represents a pre-phase for out-transition. This idea is adopted from the modeling of phone transitions in speech recognition.

For both uni-state and bi-state HMMs, self-transition

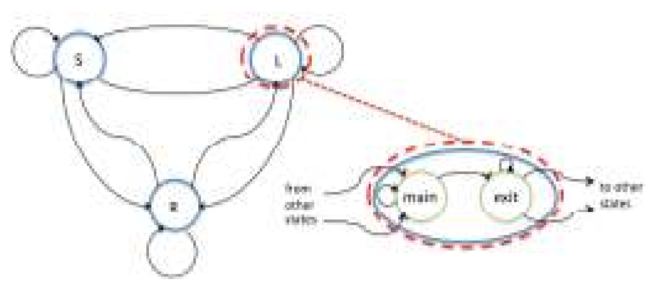

Figure 4. HMM topology with main and exit sub-states (bi-state representation). The dotted circle represents a magnified view of the state moving linearly.

probabilities are estimated from the expected motion durations; the other transition probabilities are distributed among the other allowed transitions equally. (For example in Figure 4 allowed transitions from the *main* state are self-transition and transition to the *exit* state.) Observations are modeled as Gaussian mixtures with parameters estimated from the training data based on the Maximum Likelihood principle. Initial state distribution is uniform among the three *main* sub-states of states L, R and S.

3) Temporal Classification with Discriminative Models

Conditional Random Fields are discriminative models, where the probabilities are conditioned on the observation sequences. Unlike HMMs, which depend on joint probabilities, CRFs specify the probability of possible label sequences given the observation sequence. CRFs release the independence assumption of HMMs, by allowing arbitrary relationships among the observations and dependence of hidden state probabilities on past and future observations.

Training of the CRF model was done with the hCRF library[3].

VI. Experimental Results

Evaluation of all motion classification strategies were done with 6-fold cross validation, using the dataset explained in Section IV.2. For creating a testing fold, events were picked from the whole dataset randomly, not at regular intervals. Resulting folds contained almost equal amount of events from the ideal and noisy (multiple tags and human movement in the environment) subsets. Train and test sets for cross validation were determined based on complete sequences, not feature vectors. This approach prevents the possibility that part of a sequence is used in training and remaining part of the same sequence is used for testing. We observed in our preliminary experiments that, ignoring this fact increases the classification accuracy artificially. For most of the experiments, we present moving-still classification (M/S) results in addition to the finer-grained classification results (linearly moving, randomly moving, still – $M_L/M_R/S$).

Our primary evaluation metric is the percentage of accurately predicted labels in a sequence:

$$accuracy(event) = \frac{true\ positives + true\ negatives}{total\ \#\ of\ \ labels}$$

Performance of a particular method is reported by averaging the accuracy rate over all events.

A. Effect of Interpolation

Interpolation is an extra processing for the system and it may be removed if the introduced gain is negligible. In this experiment, our aim is to evaluate the effect of interpolation

TABLE II
CLASSIFICATION RESULTS WITH AND WITHOUT INTERPOLATION

classification mode	SVM w intrp.	SVM wo intrp.	HMM w intrp.	HMM wo intrp.
M/S	91.93	86.31	93.3	69.92
$M_L/M_R/S$	78.4	66.10	81.27	58.62

[3] http://sourceforge.net/projects/hcrf/

empirically.

As demonstrated in Table II, for both classifiers and window sizes, performance was significantly better when

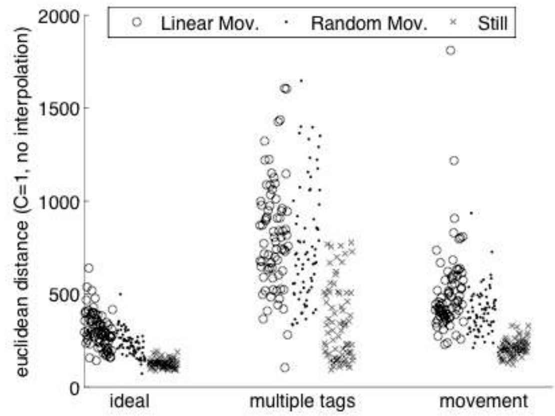

Figure 5. Euclidean distance within a window, obtained without interpolation. Environmental effects clearly affect the behavior of classes (compared to the interpolated case – Figure 3-c).

interpolation is applied. The difference in accuracy is higher for HMMs because HMM is a parametric model and a single Gaussian was not able to handle all settings (ideal, multiple tags and movement). This finding is illustrated Figure 5. For a specific movement type, the Euclidean distances computed at different settings are highly different and cannot be represented by a single parameter set. (This is not observed for standard deviation feature.) This observation necessitates using a more complicated model, such as a Gaussian mixture. However training a more complicated model requires more data to estimate the increased number of parameters. In addition, number of mixtures (corresponds to number of settings) must be known a priori. Therefore, interpolation provides a simple solution to decrease the dissimilarity in case of changing environment.

B. Effect of Classifier and Window Size

Next, we evaluate the motion monitoring performance of several classifiers and the effect of window size on classification accuracy. To implement boosting and decision tree, two binary classifiers were cascaded (will be explained in Section VI-F), whereas for SVM, HMM and CRF a single multi-class classifier was used. The overlap of two consecutive windows was adjusted to be the half-size of the window (1 second for window sizes 1.5, 2 and 3 seconds; 2 seconds for 5-second windowing; 5 seconds for 10-second windowing.

The results are depicted in Figure 6. For both classification modes (M/S and $M_L/M_R/S$), SVM demonstrated the best performance among all the non-temporal classifiers. DT and Boosting outperformed SVM only in case of the smallest window size (1.5 seconds). Accuracy with a DT was almost constant for different window sizes, especially for M/N classification mode. Therefore DTs might be advantageous for applications that have strict latency limitations.

The HMM-based classifier yielded even higher scores than SVM for both M/S and $M_L/M_R/S$ classifications. The reason is, temporal relations are better represented with an HMM. Although the SVM output is further smoothed, lost temporal information in SVM stage cannot be restored because HMM fixes transitions on the final predicted sequence only.

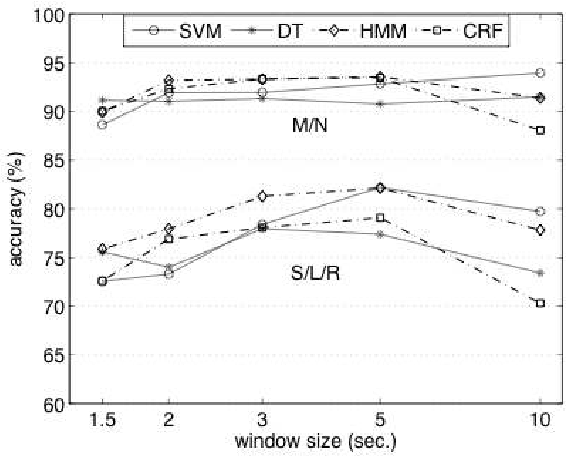

Figure 6. Comparison of classification accuracy for different classifiers and window sizes.

Accuracy scores obtained with an HMM and CRF were quite similar for M/S classification. HMM yielded better scores than HMM for finer grained ($M_L/M_R/S$) classification. Therefore, the underlying distribution was well represented with an HMM when the observations consist of the standard deviation and Euclidean distance computed within a window. We will show in Section VI-E that, a richer set of features, along with discriminative models, improves motion detection performance.

Figure 6 also shows that, a 5-second window yielded the best results for $M_L/M_R/S$ classification. No significant improvement was observed for M/N classification as the window size increased beyond 2 seconds. Therefore, a 2 second capture of RSS values was sufficient for detecting motion, however a longer capture (3 or 5 seconds) is required to identify the movement type as random or linear. In our case, a 3-second window is the optimum choice because it is only slightly worse than the 5-second window, but introduces less amount of latency.

C. Analysis on Ideal and Noisy Subsets

TABLE III
MOTION DETECTION ACCURACY ON IDEAL AND NOISY SUBSETS

window size	ideal		multiple tags + movement	
	M/S	$M_L/M_R/S$	M/S	$M_L/M_R/S$
1500	92.63	79.48	86.97	71.94
2000	94.29	77.93	91.95	78.29
3000	94.40	84.59	92.18	77.89
5000	94.76	85.87	92.27	79.11

In this section, we present the motion detection accuracy of an HMM on ideal and noisy data subsets separately. Because the events for testing were selected randomly during cross validation, it is expected that a testing fold includes almost equal amounts of ideal and noisy data, which is also valid for the training fold. For dynamic environments, such as trauma resuscitation (Section I.A), it is important to know how different a classifier performs on ideal and noisy test events.

Results are presented in Table III for several window sizes. For M/S classification, accuracy on the ideal set is about 3% higher than the accuracy of noisy subset. The difference is around 5% for $M_L/M_R/S$ classification mode. Although human movement and multiple tags in the environment reduce motion classification accuracy, the difference is not too large and can be acceptable for many applications. In a noisy environment, achievability of movement detection with 92% is an important finding that can be useful in design of real-world deployments.

We were also confronted with the cases where the noise was constructive. For instance, when window size is 2 seconds, accuracy on the noisy subset is slightly higher than the ideal subset. A similar effect can be observed in Figure 3-a, where samples from different classes are more separable in case of increased tag population, compared to the ideal setting.

D. Effect of 2-stage Classification

Both experimental results and plots of feature values (Figure 3) showed that separability of *still* from the other classes is more than the separability of *random* and *linearly moving* classes from each other. Based on this observation, it is intuitive to apply a cascaded classification strategy, where one of the binary classifiers separates moving and nonmoving events and another one for separates linear and random movements. During the testing phase, a feature sample was first classified as *moving* (linearly or randomly) or *nonmoving* (still). If the assigned label was *moving*, the second classifier was applied to identify the motion type as linear or random.

Single stage and cascaded classification results are summarized in Table IV for several window sizes and classifier types. For decision trees and boosting, the cascade of two binary classifiers yielded better scores. This is consistent with the results in [6]. However cascading did not provide any gain for SVMs. The reason is, SVM is originally a binary classifier and for multi-class classification, multiple one-to-all or pair wise classifiers are trained. In our case, we used a pair wise classifier. So, our cascading strategy was implicitly

TABLE IV
SINGLE STAGE AND CASCADED CLASSIFICATION RESULTS

window size	SVM		DT		BOOSTING	
	single	casc.	single	casc.	single	casc.
1500	72.55	70.84	74.39	75.58	58.68	72.7
2000	73.3	72.42	71.38	74.01	57.85	70.48
3000	78.4	78.82	76.68	77.95	62.72	75.15
5000	81.93	81.57	77.76	78.54	68.05	78.67
10000	82.59	82.04	78.01	78.23	66.83	77.67

embedded in the multi-class SVM.

E. *Effect of Features Individually*

In this experiment, we analyze the efficiency of features described in Section V.A.3. We experimented with several subsets.

………………

TABLE V
CLASSIFICATION RESULTS WITH VARIOUS FEATURE SETS AND CLASSIFIERS

feature set	SVM		HMM		CRF	
	M/S	M_L/M_R /S	M/S	M_L/M_R /S	M/S	M_L/M_R /S
1	84.46	65.23	82.22	62.48	83.98	61.07
2			92.73	77.71		
1,2	91.93	78.4	93.30	81.27	93.3	78.06
1,2,3,4,5	92.29	83.43	92.61	81.70	93.89	84.57

VII. CONCLUSION

……

ACKNOWLEDGMENT

…..

REFERENCES

[1] K. Muthukrishnan, M. Lijding, N. Meratnia, and P. Havinga, "Sensing motion using spectral and spatial analysis of WLAN RSSI," *Proceedings of the 2nd European conference on Smart sensing and context*, 2007, pp. 62–76.
[2] T. King and M.B. Kj\aergaard, "Composcan: adaptive scanning for efficient concurrent communications and positioning with 802.11," *Proceeding of the 6th international conference on Mobile systems, applications, and services*, 2008, pp. 67–80.
[3] M. Philipose, K.P. Fishkin, M. Perkowitz, D.J. Patterson, D. Fox, H. Kautz, and D. Hahnel, "Inferring activities from interactions with objects," *IEEE Pervasive Computing*, 2004, pp. 50–57.
[4] A. Sarcevic, I. Marsic, M.E. Lesk, and R.S. Burd, "Transactive memory in trauma resuscitation," *Proceedings of the ACM 2008 conference on Computer supported cooperative work*, 2008, pp. 215–224.
[5] D. Patterson, L. Liao, D. Fox, and H. Kautz, "Inferring high-level behavior from low-level sensors," *UbiComp 2003: Ubiquitous Computing*, 2003, pp. 73–89.
[6] T. Sohn, A. Varshavsky, A. LaMarca, M. Chen, T. Choudhury, I. Smith, S. Consolvo, J. Hightower, W. Griswold, and E. De Lara, "Mobility detection using everyday GSM traces," *UbiComp 2006: Ubiquitous Computing*, 2006, pp. 212–224.
[7] J. Krumm and E. Horvitz, "Locadio: Inferring motion and location from wi-fi signal strengths," *Proceedings of International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous' 04)*, 2004.
[8] K. Kleisouris, B. Firner, R. Howard, Y. Zhang, and R.P. Martin, "Detecting intra-room mobility with signal strength descriptors," *Proceedings of the eleventh ACM international symposium on Mobile ad hoc networking and computing*, 2010, pp. 71–80.
[9] Z. Li, C.H. Chu, and X. Wu, "Monitoring object motion using passive RFID technology," *RFID-Technology and Applications (RFID-TA), 2010 IEEE International Conference on*, 2010, pp. 233–238.
[10] B. Jiang, K.P. Fishkin, S. Roy, and M. Philipose, "Unobtrusive long-range detection of passive RFID tag motion," *Instrumentation and Measurement, IEEE Transactions on*, vol. 55, 2006, pp. 187–196.
[11] L. Ravindranath, V.N. Padmanabhan, and P. Agrawal, "Sixthsense: RFID-based enterprise intelligence," *Proceeding of the 6th international conference on Mobile systems, applications, and services*, 2008, pp. 253–266.
[12] D.J. Patterson, D. Fox, H. Kautz, and M. Philipose, "Fine-grained activity recognition by aggregating abstract object usage," *Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on*, 2005, pp. 44–51.
[13] J. Favela, M. Tentori, L.A. Castro, V.M. Gonzalez, E.B. Moran, and A.I. Martínez-García, "Activity recognition for context-aware hospital applications: issues and opportunities for the deployment of pervasive networks," *Mobile Networks and Applications*, vol. 12, 2007, pp. 155–171.
[14] B. Logan, J. Healey, M. Philipose, E.M. Tapia, and S. Intille, "A long-term evaluation of sensing modalities for activity recognition," *Proceedings of the 9th international conference on Ubiquitous computing*, 2007, pp. 483–500.
[15] R. Want, "The magic of RFID," *Queue*, vol. 2, 2004, pp. 40–48.
[16] P.V. Nikitin and K.V.S. Rao, "Antennas and propagation in UHF RFID systems," *RFID, 2008 IEEE International Conference on*, 2008, pp. 277–288.
[17] M. Buettner and D. Wetherall, "An empirical study of UHF RFID performance," *Proceedings of the 14th ACM international conference on Mobile computing and networking*, 2008, pp. 223–234.
[18] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I.H. Witten, "The WEKA data mining software: An update," *ACM SIGKDD Explorations Newsletter*, vol. 11, 2009, pp. 10–18.
[19] E. Kim, S. Helal, and D. Cook, "Human activity recognition and pattern discovery," *Pervasive Computing, IEEE*, vol. 9, 2009, pp. 48–53.
[20] T. Dietterich, "Machine learning for sequential data: A review," *Structural, Syntactic, and Statistical Pattern Recognition*, 2009, pp. 227–246.
[21] E.P.C. EPCglobal, "radio-frequency identity protocols class-1 generation-2 UHF RFID protocol for communications at 860 MHz-960 MHz version 1.0. 9," *URL: http://www. epcglobalinc. org/standards*.